ARL-TR-7317 ● JUN 2015

US Army Research Laboratory

The Effect of Resolution on Detecting Visually Salient Preattentive Features

by Adrienne Raglin and Christine Chan

NOTICES

Disclaimers

ARL-TR-7317 ● JUN 2015

US Army Research Laboratory

The Effect of Resolution on Detecting Visually Salient Preattentive Features

by Adrienne Raglin and Christine Chan
Computational and Information Sciences Directorate, ARL

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED (From - To)
Jun 2015	Final	

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
The Effect of Resolution on Detecting Visually Salient Preattentive Features	
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
Adrienne Raglin and Christine Chan	
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
US Army Research Laboratory ATTN: RDRL-CIE 2800 Powder Mill Road Adelphi, MD 20783-1138	ARL-TR-7
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Determining the factors that affect saliency within an image is valuable to a wide variety of real-world applications, such as object detection and identification. In this report, we discuss an initial investigation on the impact of 1 factor in particular, image resolution. Understanding the relationship between image resolution and saliency may lead to new image processing techniques that are able to factor in an attempt to improve the extraction of valuable information from a scene. A total of 12 scenes, 10 natural and 2 artificial, were used to create the image set used in this study, which consisted of images of each scene at 100%, 75%, 50%, 25%, and 18.75% of the original image resolution. The Harrison and Etienne-Cummings ideal observer model (IOM) was then applied to the image set so that salient areas could be highlighted. The IOM was used to identify the 20 most salient regions within each image. Then for each scene, the performance of the IOM was compared across all resolutions. We were particularly interested in whether the same salient regions persisted across all resolutions. The results of the study show that resolution may affect the selection of salient areas within the IOM.

15. SUBJECT TERMS

computational models, Ideal Observer Model, natural scenes, preattentive vision, resolution, visual salience

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Adrienne Raglin
a. REPORT	b. ABSTRACT	c. THIS PAGE	UU	16	19b. TELEPHONE NUMBER (Include area code)
Unclassified	Unclassified	Unclassified			(301) 394-0210

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Contents

List of Figures

Acknowledgments

This research was sponsored by the US Army Research Laboratory, Adelphi Laboratory Center. We would like to thank Tyrnita Moore for aiding in the pre-pilot study and the Computational and Information Sciences Directorate (CISD) students that participated in the pre-pilot study from Dr Clare Voss's team and for the insightful discussions we have had during the course of this work. We would also like to thank Andre Harrison and Ralph Etienne-Cummings for their research on the ideal observer model (IOM).

INTENTIONALLY LEFT BLANK.

1. Introduction

Whether the task is to find an individual in a packed stadium or navigate around obstacles while driving, visual searching is a key task that our visual systems perform daily. Researchers are investigating how visual systems perform tasks like these and they seek to use that knowledge to develop new image processing techniques. Feature integration theory states that there is a preattentive and attentive process to visual perception of the surrounding environment or displayed scene. "*Preattentive processing* of visual information is performed automatically on the entire visual field detecting basic features of objects in the display. Such basic features include colors, closure, line ends, contrast, tilt, curvature, and size. These simple features are extracted from the visual display in the preattentive system and later joined in the focused attention system into coherent objects. Preattentive processing is done quickly, effortlessly and in parallel without any attention being focused on the display."[1] The attentive or focused attention process combines individual features for object recognition.

Within the preattentive process, different areas of a scene are not considered equally, some areas draw an individual's attention more than others. The areas that draw more visual attention have more visually salient features. Salient features cause areas within a scene to "pop-out" or draw an individual's attention immediately (Fig. 1).[2] Visual salience is a bottom-up, stimulus-driven component of attention that is linked to the features within a scene; whereas, the top-down component is driven by the intentions and expectations of the person.

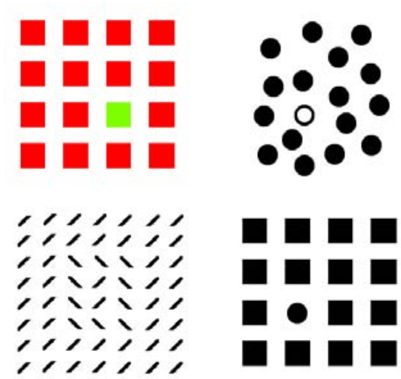

Fig. 1 Examples of salient features in artificial scenes

As effective as the preattentive process is, some situations can make this process more difficult. For example, it is more difficult to distinguish a dull yellow daffodil among a field of dull yellow dandelions versus finding a bright red rose in that same field. The human eye is directed to particular regions in a scene by highly salient features, for example, the color of the flower discussed in the previous example. These areas of interest compete for the viewer's attention as scenes become more complex. Physical differences in the visual system of 2 individuals can also lead to differences in attention. For example, if a person is colorblind, there may be differences in what draws their attention in comparison to someone who is not colorblind due to differences in what is salient between the 2 individuals. There are many other factors that influence attention such as center bias, subjective image selection, image resolution, and a person's goals.[3–5] In this study, however, we focus on the possible influences of image resolution on saliency.

Modeling visual saliency helps researchers understand and predict where a person will look within a scene. Some models attempt to replicate the physical structure of the human visual system that controls an individual's gaze. Other models create techniques based on the function and behavior of the visual system that influences what directs an individual's gaze. In general, models use low-level features such as color, intensity, and orientation to generate saliency maps. In addition, higher-order statistics have been exploited to enhance the predictive power of saliency-based models, but the extent at which they are effective still remains under investigation.[6] These types of models can be used for a wide variety of tasks, for example, navigational assistance, object recognition, and even, system design.

The model used in this investigation is a wavelet, entropy-based, saliency ideal observer model (IOM). It does not require training and relies solely on natural scene statistics. The IOM employs a bottom-up approach to select salient areas within an image.[7]

Our investigation stems from a general inquiry into resolution and its effects on preattentive salient features. For this investigation, the image resolution is varied systematically to explore its influence on identifying salient features within images. We hypothesize that resolution could be a factor in the locations that the IOM chooses as interesting or salient. Figure 2 shows an example of the saliency map used in this investigation.

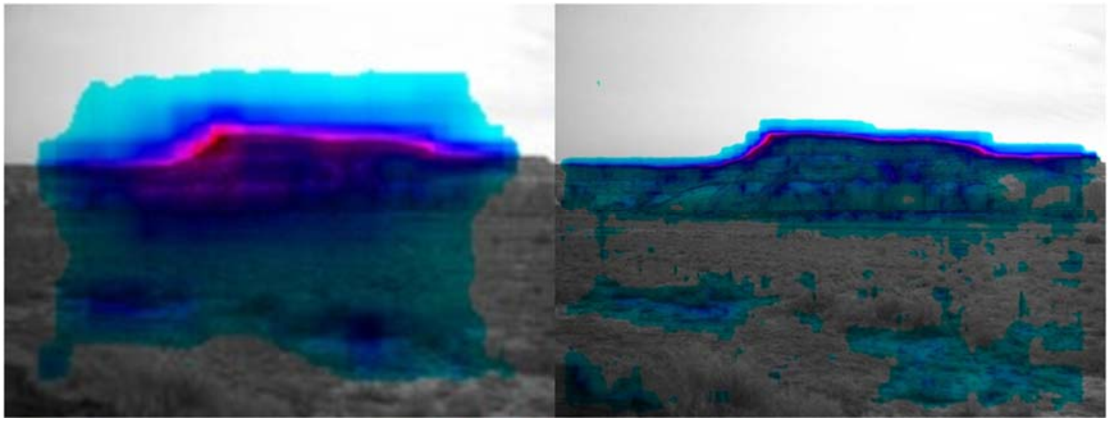

Fig. 2 A saliency map is a topographical approach to displaying ranges of saliency for an image. These saliency maps were constructed for a low (right) and high (left) resolution version of the same image. According to the IOM output, the low resolution image has a much wider diffusion of attention than the high resolution one.

2. Methods

Ten images from the Massachusetts Institute of Technology (MIT) Computer Science and Artificial Intelligence Laboratory Database of Objects and Scenes[3] were used for this pre-pilot study. These images were of unique natural scenes with a variety of compositions and subjects. Two artificial grayscale scenes were included as well. The images were then bilinearly downsampled into 4 additional resolutions: 75%, 50%, 25%, and 18.75%. The 18.75% resolution was set as the lower boundary due to the IOM minimum size limitation for input images. The bilinear downsample was done is such a way that aliasing was not present in the final image. A data set comprised of the top 20 areas of interest that the IOM selected from each image resolution was obtained. Sixty data sets were obtained and prepared for analysis by recording the coordinates for the salient areas of each image in a spreadsheet to aide in the calculations. The salient locations for each resolution were rescaled to the dimensions of the 50% resolution image in order to provide a common resolution. The rescaling linearly expands or shrinks the pixel locations of salient areas to match the dimensions of the 50% resolution image. Scatter plots provide a visual representation of the coordinates of the salient areas for each of the 5 image resolutions (Fig. 3).

A cluster is defined as a group of points that share a coordinate set with a deviation limit of ±5. Efforts were made to review each plot to determine potential issues with the data collected. Salient areas that were different for the same image set or similar across different images were highlighted. If an error occurred using these data sets in the IOM, the resolution was checked and the procedure repeated.

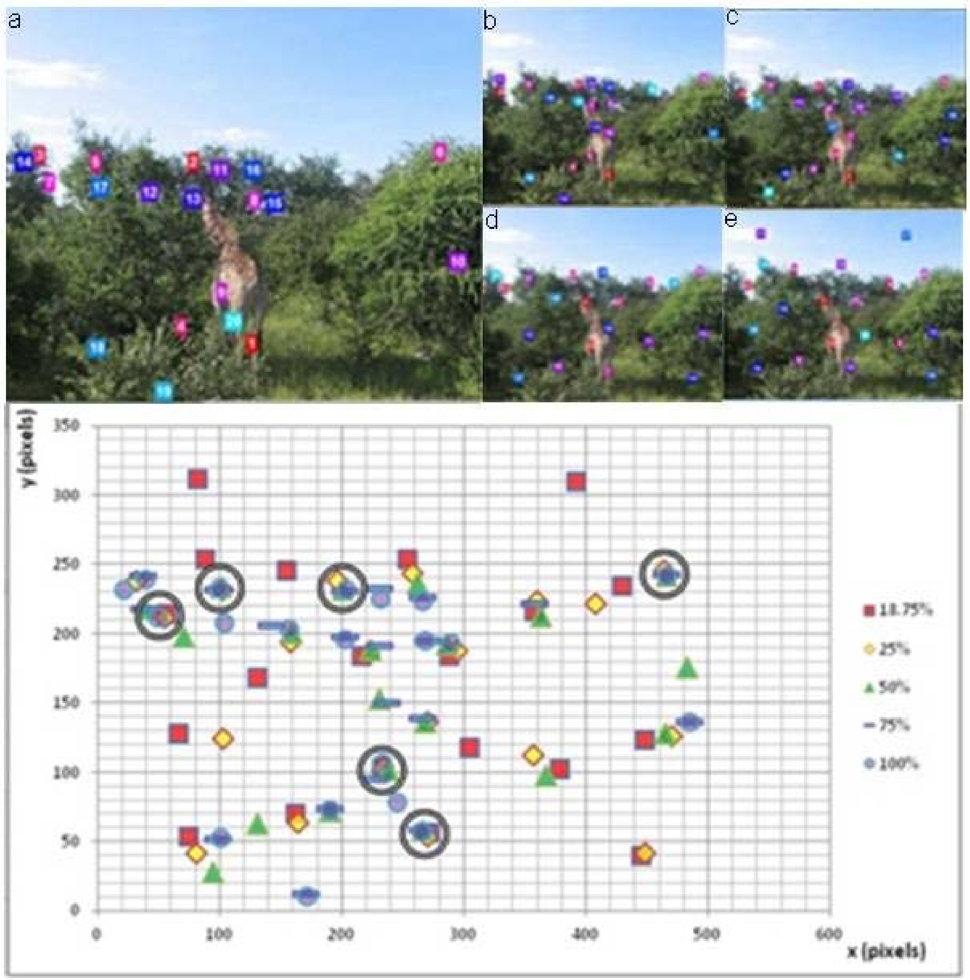

Fig. 3 Sample image shown in 4 resolutions in descending order (a–e). The plot compiles the areas of interest displayed in the images and each symbol represents 1 of the images. Data clusters indicate where the IOM identified areas of interest across all resolutions. The circles were superimposed over the data clusters for emphasis.

3. Results and Discussion

When the salient areas highlighted by the IOM were overlaid onto the corresponding images, the results showed a number of clusters as well as non-clustered areas. One of the first steps of the IOM is decomposing the source image into smaller samples of the image. Therefore, inherently, the IOM does not process an image differently at varying resolutions. However, high resolutions images can be downsampled more. The results indicate that because of the lower resolution the IOM processed images differently. A question for future investigation would be exploring the decomposition of the original image to determine if other factors were involved. Had resolution been a non-factor of IOM-replicated saliency, the graphs should show overlapping clusters. The number of clusters varied among image sets. The lowest resolution data sets all displayed a lattice-like pattern in identified areas of interest, whereas the higher resolution data sets showed areas of interest more closely concentrated around salient areas. The artificial controls showed a similar effect to the natural images.

The study results suggest that resolution is an influential factor. Major observations included salient areas that did not cluster as resolution changed and salient areas appeared in uniformed patterns as the resolution of the image decreased. It is unclear if the treatment of neighboring pixels or processing of images of varying size might contribute to the lattice like pattern of the results from the lower resolution data sets. Also, the different percentages of the resolution selected may have influenced the effect of resolution change in the IOM.

4. Conclusion and Future Work

Given the variation within the data sets, these initial results indicate there may be a shift between the areas of interest identified by the IOM as an image decreases in resolution. As resolutions decrease, IOM-identified areas of interest appeared to progressively change from highly salient areas found at higher resolutions; salient areas in the higher resolution images seemed more likely to cluster. However, other features may play more influential roles than changes in resolution.

The findings suggest that resolution may have an effect on overall preattention simulation in the IOM. The results discussed were a part of an initial pre-pilot study only; additional studies would be needed to further investigate the trends found in these results, such as the inconsistencies in the clusters throughout the images sets that may indicate underlying confounding variables. Also, additional investigation can explore the relationship between clustering of salient areas and the rank the IOM places on the salient areas.

Further exploration of the idea initiated in this study may reveal new information concerning the degree of influence resolution has on traditional preattentive features and potentially improve image processing techniques. Additional research using larger quantities of images with a wider variety of compositions to investigate the relationship between resolution and other salient features would be recommended. A practical next step would be to differentiate between size and pixel density as an added definition to the resolution factor. To more thoroughly test this proposition, comparative studies of resolution variations in other saliency models as well may be conducted. This may provide further evidence to support resolution influence if other models display trends similar to the IOM's. Including ground truth of human visual data from studies where saccade positions were measured when subjects were presented imagery with varying resolutions may also aid further research.

5. References

1. Treisman A. Features and objects in visual processing. Scientific American. 1986;255(5):114–125.

2. Kadir T, Brady M. Saliency, scale and image description. Oxford: University of Oxford, 2000.

3. Bruce N, Tsotsos J. Saliency, attention, and visual search: an information theoretic approach. Journal of Vision. 2009;9(3):1–24.

4. Harrison A, Etienne-Cummings R. A Entropy based Saliency model using Natural Scene Statistics. Baltimore: Johns Hopkins University, 2012.

5. Tseng P, Carmi R, Camero, I, Munoz D, Itti L. Quantifying center bias of observers in free viewing of dynamic natural scenes. Journal of Vision. 2009;9(7):1–16.

6. Zhao Q, Koch C. Learning a saliency map using fixated locations in natural scenes. Journal of Vision. 2011;11(3):1–15.

7. Scenes and objects database. Cambridge (MA): Massachusetts Institute of Technology; (n.d.) [accessed January 15, 2013]. http://web.mit.edu/torralba/www/database.html.

1 (PDF)	DEFENSE TECHNICAL INFORMATION CTR DTIC OCA
2 (PDF)	DIRECTOR US ARMY RSRCH LAB RDRL CIO LL RDRL IMAL HRA RECORDS MGMT
1 (PDF)	GOVT PRNTG OFC A MALHOTRA
1 (PDF)	DIRECTOR US ARMY RSRCH LAB RDRL CIE A RAGLIN

INTENTIONALLY LEFT BLANK.